ABSTRACT OF THE DISCLOSURE

A memory system having a number of partitions each operative to independently service memory requests from a plurality of memory clients while maintaining the appearance to the memory client of a single partition memory subsystem. The memory request specifies a location in the memory system and a transfer size. A partition receives input from an arbiter circuit which, in turn, receives input from a number of client queues for the partition. The arbiter circuit selects a client queue based on a priority policy such as round robin or least recently used or a static or dynamic policy. A router receives a memory request, determines the one or more partitions needed to service the request and stores the request in the client queues for the servicing partitions. In one embodiment, an additional arbiter circuit selects memory requests from one of a subset of the memory clients and forwards the requests to a routing circuit, thereby providing a way for the subset of memory clients to share the client queues and routing circuit. Alternatively, a memory client can make requests directed to a particular partition in which case no routing circuit is required. For a read request that requires more than one partition to service, the memory system must collect the read data from read queues for the various partitions and deliver the collected data back to the proper client. Read queues can provide data in non-fifo order to satisfy an memory client that can receive data out-of-order.